MessageLabs
the message is clear™

The Skeptic Factor

How Skeptic™, MessageLabs' proprietary anti-virus technology, works at the Internet level to instantly identify new outbreaks

When the Goner virus hit the business world in December 2001, MessageLabs' SkyScan Anti-Virus service identified it immediately. Goner was one of the largest virus outbreaks of recent times, causing major business disruption internationally, yet MessageLabs' customers were completely unaffected.

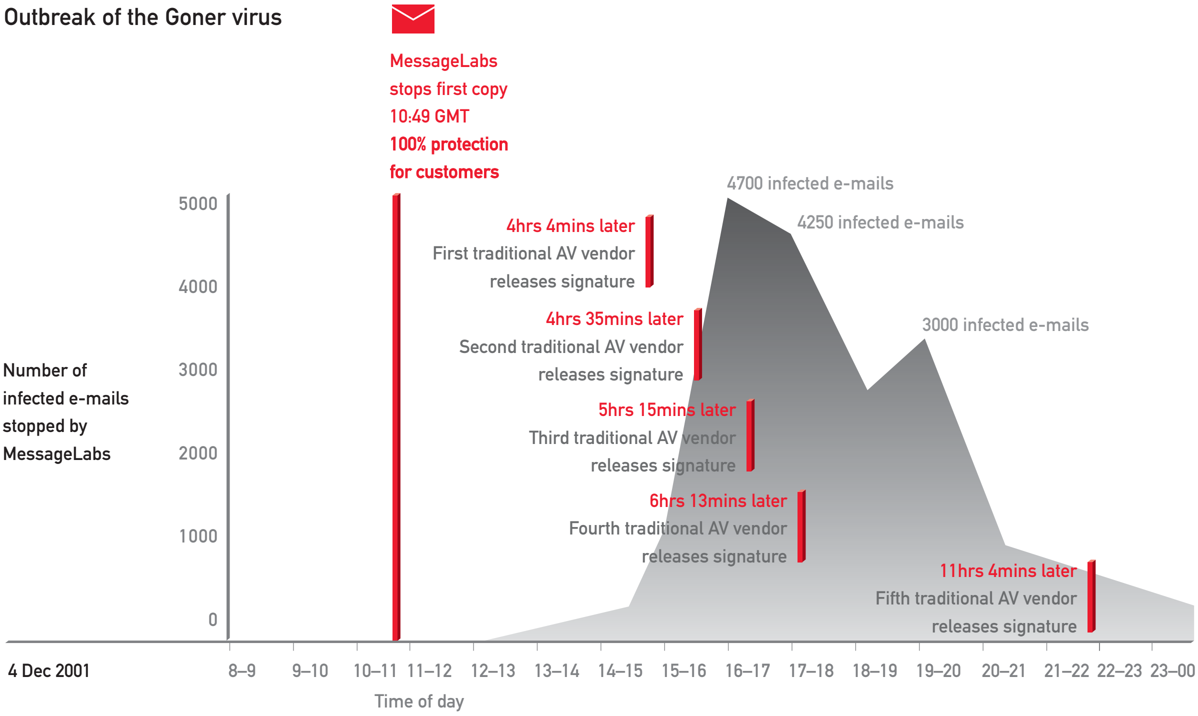
Meanwhile, the rest of the world was in chaos, either cleaning up systems or taking preventative action in shutting down e-mail servers and waiting for anti-virus software vendors to provide a fix. As usual in a new outbreak, fixes did not become available for several hours.

Similarly, SkyScan AV instantly identified the infamous LoveBug when it struck in May 2000. Again, it took several hours for the traditional software vendors to provide new updates. These then needed to be distributed to millions of users around the world, with multiple platforms, locations and time zones.

It has been the same story every time a new outbreak has occurred. Major virus incidents such as Kournikova, Homepages, SirCam, Nimda, BadTrans and Klez have all caused havoc in the business world – but not for MessageLabs' customers.

So, why is it that MessageLabs can immediately identify a new virus the moment it appears – but traditional software vendors can only leave their customers floundering without e-mail until such time as a signature is created and distributed?

Outbreak of the Goner virus



The answer is Skeptic™

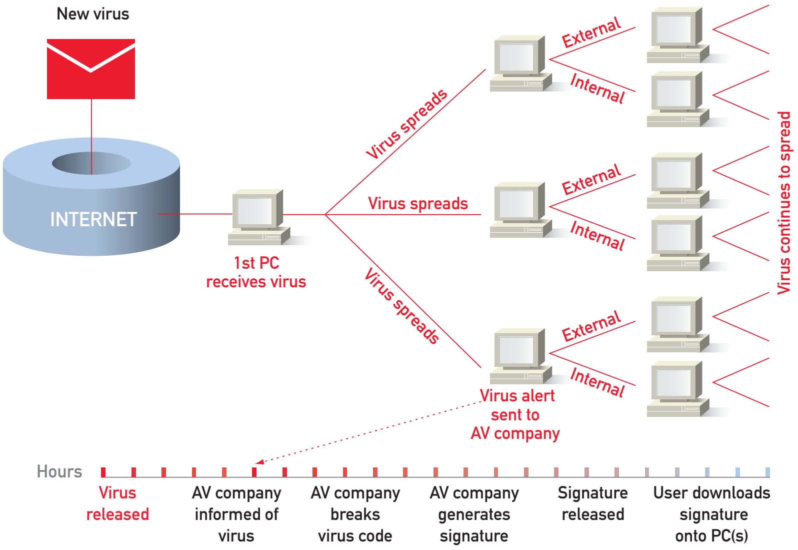
Skeptic is MessageLabs' proprietary anti-virus technology, designed to identify new viruses without any need for signature updates. From its unique position, operating at the Internet level, rather than server or desktop level, Skeptic can identify techniques or characteristics which are indicative of an e-mail virus. These include movement patterns, forwarding e-mail with the same or similar details to multiple names in the intended recipient's address book.

Thus, Skeptic can identify a virus even if it is completely original in construction – and our customers are automatically protected from it.

Compare this instant proactive capability with the reactive methods of conventional anti-virus software. Traditional software is dependent on at least one customer recognising it has been infected or to become suspicious about a particular file.

The customer then needs to send a sample in for analysis which will eventually lead to a fix being produced ready for distribution. This all takes time – while the virus keeps on spreading. However, the unique architecture of the MessageLabs solution ensures that the virus is proactively stopped at the Internet level – before it comes anywhere near a customer's network boundary.

How traditional anti-virus software works



New virus
INTERNET
1st PC receives virus
Virus spreads
Virus spreads
Virus spreads
External
Internal
External
Internal
External
Internal
Virus continues to spread
Virus alert sent to AV company
Hours
Virus released
AV company informed of virus
AV company breaks virus code
AV company generates signature
Signature released
User downloads signature onto PC(s)

So, how does Skeptic work?

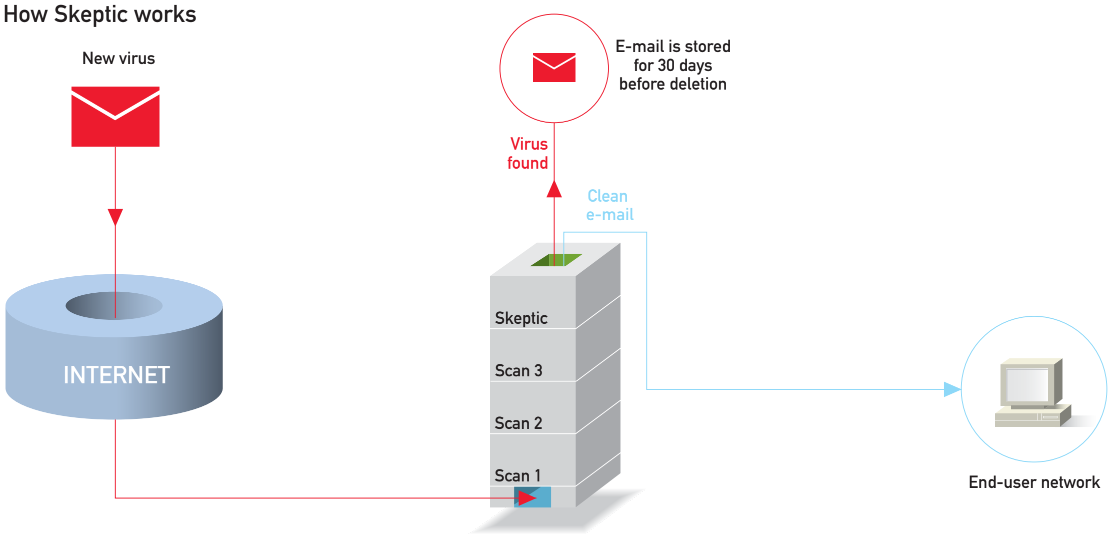
Traditional reactive anti-virus software, dependent on new signature updates being available and uploaded, is effective at stopping the back catalogue of viruses, but its weaknesses are exposed during a new outbreak.

Skeptic, on the other hand, uses heuristics methods developed explicitly to identify new viruses in e-mail – at the Internet level. These are different from the heuristics methods used by traditional desktop and gateway anti-virus products and have proven themselves time and again by stopping all of the major outbreaks of recent times.

Many anti-virus companies have invested in desktop or gateway heuristics, designed to stop new viruses, but the large number of outbreaks which we have seen over the last few years shows that their technology still has some way to go.

Skeptic was designed to address these weaknesses and draw on the huge e-mail volumes and global visibility which the MessageLabs network provides.
Of course, Skeptic is also completely different in that it applies heuristics at the Internet – rather than server or desktop – level, long before any infected e-mail can get into the customer's internal networks.

How Skeptic works



Adaptive heuristics

We coined the term 'adaptive heuristics' to describe how Skeptic applies heuristics technology within a global architecture. Skeptic operates as part of a managed service at the Internet level, giving it a number of key advantages over individual programs.

The key advantage is that Skeptic has visibility of the millions of e-mails and thousands of viruses passing through the MessageLabs network every day, providing a complete and immediate view of any potential threat.

This allows Skeptic to build a detailed profile of what both good and bad e-mail looks like, and global database replication architecture allows knowledge gained in one part of our network to be replicated globally within seconds. This technology is simply not practical in the desktop environment and is one of the key differentiators between desktop and Internet-level heuristics.

A standard global platform

Skeptic works in real time on a common platform and can, therefore, be much more effective than heuristics which first have to be tested on multiple platforms and distributed to millions of users in locations across the world.

An average anti-virus program is about 2Mb in size, because of the processing power of the machines on which it is designed to run, leading to severe limitations in the levels of sophistication within the software. Skeptic runs on high-powered, globally distributed architecture and has a 2Gb knowledge base. This means that it is capable of far greater performance and functionality than would ever be possible within client- or server-based software.

Knowledge to the power of 7 million

Since the introduction of the SkyScan Anti-Virus service in 1999, the MessageLabs network has intercepted more than 7 million viruses. Each one of them has automatically been broken down into its component parts and entered into Skeptic's knowledge base of virus-like behaviour and techniques. Because most new viruses are based at least partly on viruses which have gone before, this enables Skeptic to identify and stop new viruses.

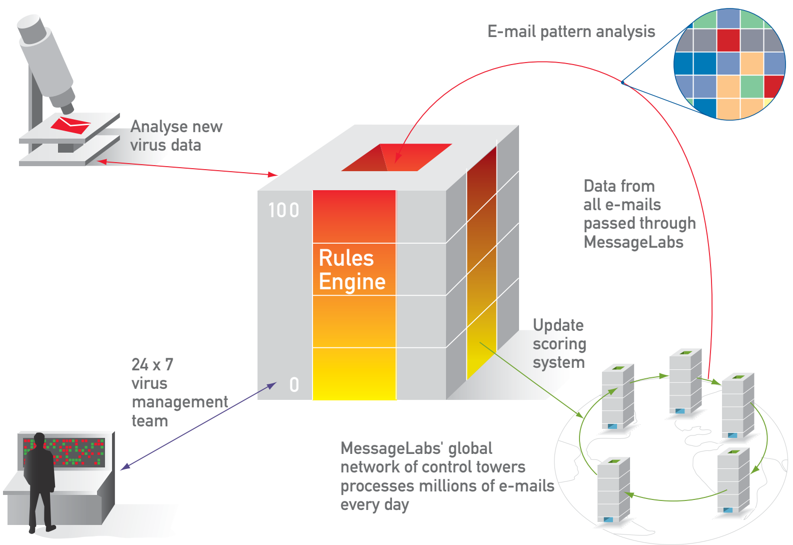
For example, Skeptic's code analysis data store has detailed profiles of virus-like behaviour for activities, including:

- Duplication – attempts to duplicate by techniques such as writing to executable files or mass-mailing
- Payload – signs of a payload-like mass deletion or unusual trigger conditions
- Trojans – signs of keyboard logging or mailing of sensitive files
- Obfuscation – ways of hiding from traditional desktop scanners

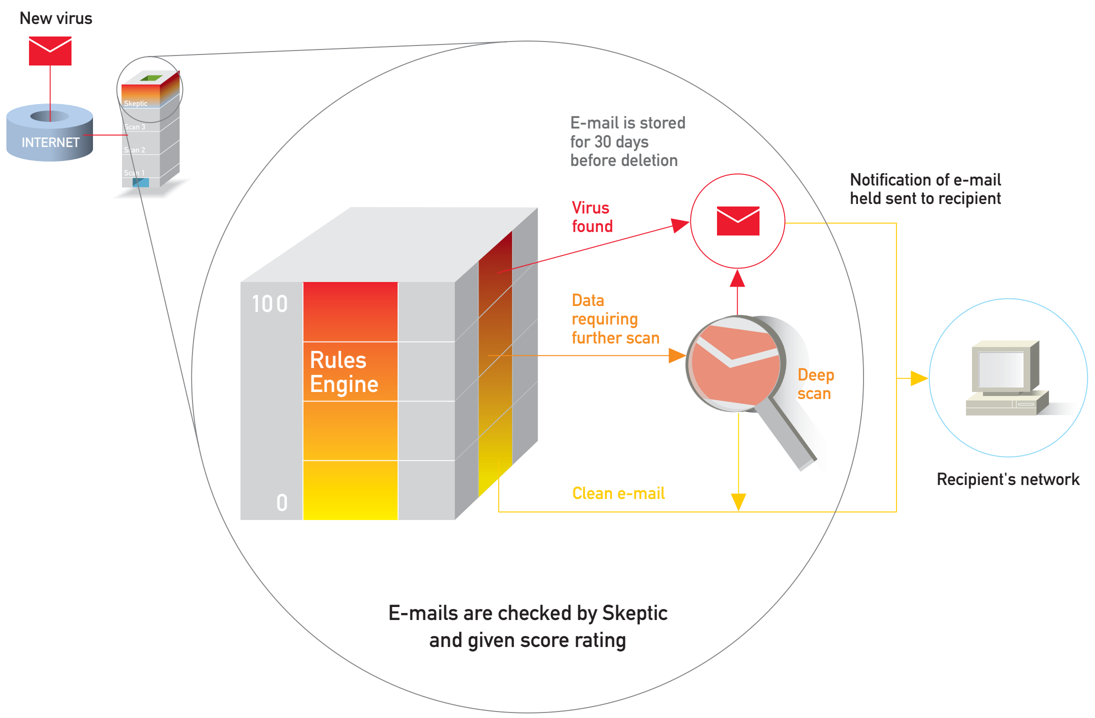
The data store knows about normal program behaviour which is similar to these activities and so can eliminate these from further consideration. It also includes a detailed library of typical tricks and exploits used by viruses, such as:

- Filenames – double extensions, hidden extensions and exploitable filenames
- Auto-executing – e-mails which cause code to execute automatically on opening
- Buffer overruns – exploitable buffer overruns in software packages, especially including e-mail software
- E-mail exploits – exploits which make use of known holes in e-mail products
- File-type-matching – Skeptic has detailed checks to ensure that file content matches the filename, which go well beyond simple file-extension-checking
- Movement patterns – a mass-mailing virus will follow specific patterns for maximum distribution

How the Skeptic data store works


E-mail pattern analysis
Analyse new virus data
100
Rules Engine
0
Data from all e-mails passed through MessageLabs
Update scoring system
24 x 7 virus management team
MessageLabs' global network of control towers processes millions of e-mails every day

How the Skeptic points system works


New virus
INTERNET
Skeptic
Scan 3
Scan 2
Scan 1
E-mail is stored for 30 days before deletion
Notification of e-mail held sent to recipient
Virus found
100
Rules Engine
0
Data requiring further scan
Deep scan
Clean e-mail
Recipient's network
E-mails are checked by Skeptic and given score rating

The points system

Skeptic uses a system in which each suspicious aspect of an e-mail is given a points rating. If an e-mail exceeds the threshold score, it is automatically stopped and quarantined as viral. Reaching a lower score will cause deeper levels of analysis to kick in.

These levels of analysis are not possible for traditional anti-virus software, without slowing the system to unacceptable levels.

Minimising false positives

Skeptic keeps false positives to a minimum by the use of trial heuristics. New heuristics are first added in logging mode – this just keeps count of the e-mails which would have been stopped if the heuristic were live. After a period of running in trial mode, results are analysed and the heuristic either discarded or accepted. A genetic algorithm is then used to assign the points score to the heuristic; constructs found only in malware will get a high score, while constructs also found in legitimate e-mail will be assigned a lower score.

Since we process more than 7 million e-mails each day, to and from companies all around the world, the learning process is both rapid and comprehensive. It also means that speed and accuracy improve. As the MessageLabs customer base increases, we process more e-mail and more diverse e-mail. Currently, our false positive ratio is around one in every one million – hardly a major problem, but we're still committed to improving it.

A record which speaks for itself

Obviously, we cannot reveal the exact algorithms used by Skeptic's heuristics technology – virus-writers would use such information to try to defeat the heuristics.

However, we can point to our record. Since the SkyScan AV service went live in 1999, MessageLabs has consistently maintained the most successful track record of any anti-virus company.

Based on current statistics, the chance of getting infected with an e-mail virus is at least 10^5 times less for a company using MessageLabs than for a company using a traditional desktop/gateway/server-only solution.

The future

Skeptic will continue to be an instant reflection of the knowledge learned from the millions of e-mails passing through the MessageLabs global network every day. As the number of e-mails being scanned continues to grow, so Skeptic will continue proactively to identify and stop new viruses, both now and in the future.

To find out more about Skeptic and other MessageLabs services, visit our Web site at www.messagelabs.com or e-mail: info@messagelabs.com